



(12) 发明专利申请

(10) 申请公布号 CN 104503947 A

(43) 申请公布日 2015. 04. 08

(21) 申请号 201410781480. X

(22) 申请日 2014. 12. 16

(71) 申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为
总部办公楼

(72) 发明人 程龙飞

(74) 专利代理机构 北京弘权知识产权代理事务
所(普通合伙) 11363

代理人 逯长明 许伟群

(51) Int. Cl.

G06F 15/163(2006. 01)

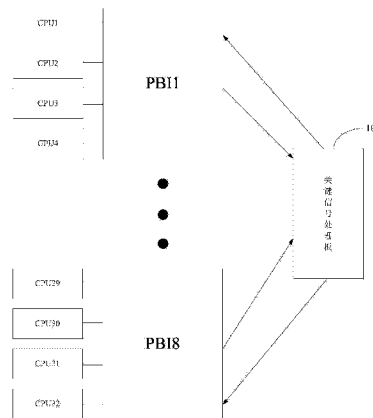
权利要求书3页 说明书12页 附图2页

(54) 发明名称

多路服务器及其信号处理方法

(57) 摘要

本发明公开了一种多路服务器及其信号处理方法,所述多路服务器包括多个CPU、多个PBI板,以及关键信号处理板,每个PBI板监控至少两个CPU。各个PBI板的输出端分别连接关键信号处理板的输入端,各个PBI板的输入端分别连接关键信号处理板的输出端。所述多路服务器通过关键信号处理板接收各个PBI板发出的关键信号,然后,同时向相应的PBI板发送与所述关键信号对应的回应信号,从而保证CPU收到的回应信号满足多路服务器的要求。由上述内容可知,各个PBI板直接与关键信号处理板连接,不需要切换开关进行串联,减少切换开关的数量,结构简单、可靠性高。而且减少了信号所经过的切换开关的数量,提高了信号质量。



1. 一种多路服务器,其特征在于,包括:多个 CPU、多个南桥基板管理监控器输入输出 PBI 板,以及关键信号处理板,其中,一个 PBI 板监控至少两个 CPU;

所述 PBI 板连接自身监控的 CPU;

各个 PBI 板的输出端连接所述关键信号处理板的输入端,各个 PBI 板的输入端连接所述关键信号处理板的输出端,所述关键信号处理板接收所述 PBI 板发送的关键信号,并向 PBI 板返回与所述关键信号相对应的回应信号;

接收到所述回应信号的 PBI 板将所述回应信号分别发送给自身监控的 CPU。

2. 根据权利要求 1 所述的多路服务器,其特征在于,所述关键信号处理板接收所述 PBI 板发送的关键信号,并向所述 PBI 板返回与所述关键信号相对应的回应信号,包括:

当所述多路服务器作为一个服务器进行上电时,所述关键信号处理板接收到所述多路服务器内的全部 PBI 板分别发送的第一类电源上电信号后,同时向所述全部 PBI 板发送作为与所述第一类电源上电信号相对应的回应信号的第二类电源上电信号,以使所述多路服务器内的全部 CPU 接收到第二类电源上电信号的时间差满足预设时间要求;

若所述多路服务器内的全部 CPU 划分成至少两个子服务器,当任意一个子服务器上电时,所述关键信号处理板接收到所述子服务器内的全部 PBI 板分别发送的所述第三类电源上电信号后,同时向所述子服务器内的全部 PBI 板分别发送作为与所述第三类电源上电信号相对应的第四类电源上电信号,以使所述子服务器内的全部 CPU 接收到所述第四类电源上电信号的时间差满足预设时间要求;

其中,所述第一类电源上电信号和所述第三类电源上电信号均由所述 PBI 板检测到自身监控的各个 CPU 的电源均上电完成时产生。

3. 根据权利要求 1 所述的多路服务器,其特征在于,所述关键信号处理板接收所述 PBI 板发送的关键信号,并向所述 PBI 板返回与所述关键信号相对应的回应信号,包括:

当所述多路服务器中插入新的 CPU 时,所述关键信号处理板接收到所述新的 CPU 对应的 PBI 板发送的第五类电源上电信号后,在预设时刻向所述新的 CPU 对应的 PBI 板发送作为与所述第五类电源上电信号相对应的回应信号的第六类电源上电信号,以使所述新的 CPU 接收到所述第六类电源上电信号的时间满足预设时间要求;

所述第五类电源上电信号由所述 PBI 板检测到所述新的 CPU 上电完成时产生。

4. 根据权利要求 1 所述的多路服务器,其特征在于,所述关键信号处理板接收所述 PBI 板发送的关键信号,并向所述 PBI 板返回与所述关键信号相对应的回应信号,包括:

若所述多路服务器作为一个服务器使用,所述关键信号处理板接收到任意一个 PBI 板发送的第一系统错误信号后,分别向所述多路服务器中除发送所述系统错误信号 PBI 之外的其它各个 PBI 板发送作为与所述第一系统错误信号对应的回应信号的第二系统错误信号;

若所述多路服务器划分成至少两个子服务器,所述关键信号处理板接收到任意一个 PBI 板发送的第三系统错误信号时,根据获得的所述子服务器信息确定与发送所述第三系统错误信号的 PBI 板属于同一个子服务器的其它各个 PBI 板;分别向所述子服务器中除发送所述第三系统错误信号的 PBI 板之外的其它各个 PBI 板发送作为与所述第三系统错误信号对应的回应信号的第四系统错误信号。

5. 根据权利要求 4 所述的多路服务器,其特征在于,所述关键信号处理板接收所述 PBI

板发送的关键信号,并向所述 PBI 板返回与所述关键信号相对应的回应信号,还包括:

当检测到当前 PBI 板发送的第一系统错误信号或第三系统错误由有效状态变为无效状态时,所述关键信号处理板接收下一个第一系统错误信号或第三系统错误信号。

6. 一种多路服务器的信号处理方法,其特征在于,应用于多路服务器中,所述多路服务器包括多个 CPU、多个南桥基板管理监控器输入输出 PBI 板,以及关键信号处理板,其中,一个 PBI 板监控至少两个 CPU;所述方法包括:

所述关键信号处理板接收所述 PBI 板发送的关键信号;

所述关键信号处理板根据接收到的所述关键信号向所述 PBI 板返回与所述关键信号相对应的回应信号,以使接收到所述回应信号的所述 PBI 板将所述回应信号发送给所述 PBI 板监控的 CPU。

7. 根据权利要求 6 所述的方法,其特征在于,所述关键信号处理板根据接收到的所述关键信号向 PBI 板返回与所述关键信号相对应的回应信号,包括:

当所述多路服务器作为一个服务器进行上电时,所述关键信号处理板接收到所述多路服务器内的全部 PBI 板分别发送的第一类电源上电信号后,同时向全部 PBI 板发送作为与所述第一类电源上电信号相对应的回应信号的第二类电源上电信号,以使所述多路服务器内的全部 CPU 接收到第二类电源上电信号的时间差满足预设时间要求;

若所述多路服务器内的全部 CPU 划分成至少两个子服务器,当任意一个子服务器上电时,所述关键信号处理板接收到所述子服务器内的全部 PBI 板分别发送的所述第三类电源上电信号后,同时向所述子服务器内的全部 PBI 板分别发送作为与所述第三类电源上电信号相对应的回应信号的第四类电源上电信号,以使所述子服务器内的全部 CPU 接收到所述第四类电源上电信号的时间差满足预设时间要求;

其中,所述第一类电源上电信号和所述第三类电源上电信号均由所述 PBI 板检测到自身监控的各个 CPU 的电源均上电完成时产生。

8. 根据权利要求 6 所述的方法,其特征在于,所述关键信号处理板根据接收到的所述关键信号向 PBI 板返回与所述关键信号相对应的回应信号,包括:

当所述多路服务器中插入新的 CPU 时,所述关键信号处理板接收到所述新的 CPU 对应的 PBI 板发送的第五类电源上电信号后,在预设时刻向所述新的 CPU 对应的 PBI 板发送作为与所述第五类电源上电信号相对应的回应信号的第六类电源上电信号,以使所述新的 CPU 接收到所述第六类电源上电信号的时间满足预设时间要求;

所述第五类电源上电信号由所述 PBI 板检测到所述新的 CPU 上电完成时产生。

9. 根据权利要求 8 所述的方法,其特征在于,所述关键信号处理板根据接收到的所述关键信号向 PBI 板返回与所述关键信号相对应的回应信号,包括:

若所述多路服务器作为一个服务器使用,所述关键信号处理板接收到任意一个 PBI 板发送的第一系统错误信号后,同时向所述多路服务器中除发送所述系统错误信号 PBI 之外的其它各个 PBI 板发送作为与所述第一系统错误信号对应的回应信号的第二系统错误信号;

若所述多路服务器划分成至少两个子服务器,所述关键信号处理板获取所述多路服务器的子服务器信息,当接收到任意一个 PBI 板发送的第三系统错误信号时,根据所述子服务器信息确定与发送所述第三系统错误信号的 PBI 板属于同一个子服务器的其它各个 PBI

板 ;同时向所述子服务器中除发送所述第三系统错误信号的 PBI 板之外的其它各个 PBI 板发送作为与所述第三系统错误信号对应的回应信号的第四系统错误信号。

10. 根据权利要求 9 所述的方法,其特征在于,所述方法还包括:

当检测到当前 PBI 板发送的第一系统错误信号或第三系统错误由有效状态变为无效状态时,所述关键信号处理板接收下一个第一系统错误信号或第三系统错误信号。

多路服务器及其信号处理方法

技术领域

[0001] 本发明涉及服务器技术领域,特别是涉及一种多路服务器及其信号处理方法。

背景技术

[0002] 多路服务器是指服务器内物理 CPU 的数量有多个,服务器中物理 CPU 的数量决定服务器中“多路”的数量,例如,服务器中包含 16 个物理 CPU,则为 16 路服务器。多路服务器中物理 CPU 的数量越多,其处理数据的能力越高,因此,16 路和 32 路的服务器成为多路服务器的发展趋势。

[0003] 若实现 32 路服务器,必须实现 32 路内信号全互联,且对时序要求极为严格。例如,在多路服务器开机时,所有的 CPU 收到表示电源正常的 Powergood 信号的时间差必须监控在 10ns 以内,并记录该时刻为 T0;在实现 CPU 热插拔、online 或 offline 时,新上线的 CPU(Intel Ivybridge 处理器)收到 Powergood 的信号时间必须为 T0+864bc1ks(若采用 Intel Haswell 处理器,则为 T0+384bc1ks)。而 Intel 公司只给出了 8 路及以下的服务器中关键信号(例如,Caterr、PMSYNC、Powergood、TSC 等信号)的处理方式,对于 8 路以上的服务器没有给出任何推荐方案。

[0004] 假设采用 Intel 给出的 8 路服务器对关键信号的处理方式实现 16 路及以上的服务器,由于 8 路服务器是通过切换开关将不同基本单板(每个基本单板上设置有多个 CPU,所述单板上的 CPU 数量是最小硬分区的单位)串联在一起,需要切换开关数量多,当传递关键信号时,需要控制相应的切换开关联动才能实现,实现过程极其复杂、可靠性差,且经过多级带有内阻的切换开关后,信号质量变得很差。

发明内容

[0005] 本发明实施例中提供了一种多路服务器及其信号处理方法,以解决现有技术中的方案复杂、可靠性差集信号质量差的问题。

[0006] 为了解决上述技术问题,本发明实施例公开了如下技术方案:

[0007] 第一方面,本发明提供一种多路服务器,包括:多个 CPU、多个南桥基板管理监控器输入输出 PBI 板,以及关键信号处理板,其中,一个 PBI 板监控至少两个 CPU;

[0008] 所述 PBI 板连接自身监控的 CPU;

[0009] 各个 PBI 板的输出端连接所述关键信号处理板的输入端,各个 PBI 板的输入端连接所述关键信号处理板的输出端,所述关键信号处理板接收所述 PBI 板发送的关键信号,并向所述 PBI 板返回与所述关键信号相对应的回应信号;

[0010] 接收到所述回应信号的 PBI 板将所述回应信号分别发送给自身监控的 CPU。

[0011] 结合第一方面,在第一方面的第一种可能的实现方式中,所述关键信号处理板接收所述 PBI 板发送的关键信号,并向所述 PBI 板返回与所述关键信号相对应的回应信号,包括:

[0012] 当所述多路服务器作为一个服务器进行上电时,所述关键信号处理板接收到所述

多路服务器内的全部 PBI 板分别发送的第一类电源上电信号后,同时向所述全部 PBI 板发送作为与所述第一类电源上电信号相对应的回应信号的第二类电源上电信号,以使所述多路服务器内的全部 CPU 接收到第二类电源上电信号的时间差满足预设时间要求;

[0013] 若所述多路服务器内的全部 CPU 划分成至少两个子服务器,当任意一个子服务器上电时,所述关键信号处理板接收到所述子服务器内的全部 PBI 板分别发送的所述第三类电源上电信号后,同时向所述子服务器内的全部 PBI 板分别发送作为与所述第三类电源上电信号相对应的第四类电源上电信号,以使所述子服务器内的全部 CPU 接收到所述第四类电源上电信号的时间差满足预设时间要求;

[0014] 其中,所述第一类电源上电信号和所述第三类电源上电信号均由所述 PBI 板检测到自身监控的各个 CPU 的电源均上电完成时产生。

[0015] 结合第一方面,在第一方面的第二种可能的实现方式中,所述关键信号处理板接收所述 PBI 板发送的关键信号,并向所述 PBI 板返回与所述关键信号相对应的回应信号,包括:

[0016] 当所述多路服务器中插入新的 CPU 时,所述关键信号处理板接收到所述新的 CPU 对应的 PBI 板发送的第五类电源上电信号后,在预设时刻向所述新的 CPU 对应的 PBI 板发送作为与所述第五类电源上电信号相对应的回应信号的第六类电源上电信号,以使所述新的 CPU 接收到所述第六类电源上电信号的时间满足预设时间要求;

[0017] 所述第五类电源上电信号由所述 PBI 板检测到所述新的 CPU 上电完成时产生。

[0018] 结合第一方面,在第一方面的第三种可能的实现方式中,所述关键信号处理板接收所述 PBI 板发送的关键信号,并向所述 PBI 板返回与所述关键信号相对应的回应信号,包括:

[0019] 若所述多路服务器作为一个服务器使用,所述关键信号处理板接收到任意一个 PBI 板发送的第一系统错误信号后,分别向所述多路服务器中除发送所述系统错误信号 PBI 之外的其它各个 PBI 板发送作为与所述第一系统错误信号对应的回应信号的第二系统错误信号;

[0020] 若所述多路服务器划分成至少两个子服务器,所述关键信号处理板接收到任意一个 PBI 板发送的第三系统错误信号时,根据获得的所述子服务器信息确定与发送所述第三系统错误信号的 PBI 板属于同一个子服务器的其它各个 PBI 板;分别向所述子服务器中除发送所述第三系统错误信号的 PBI 板之外的其它各个 PBI 板发送作为与所述第三系统错误信号对应的回应信号的第四系统错误信号。

[0021] 结合第一方面的第三种可能的实现方式,在第一方面的第四种可能的实现方式,所述关键信号处理板接收所述 PBI 板发送的关键信号,并向所述 PBI 板返回与所述关键信号相对应的回应信号,还包括:

[0022] 当检测到当前 PBI 板发送的第一系统错误信号或第三系统错误由有效状态变为无效状态时,所述关键信号处理板接收下一个第一系统错误信号或第三系统错误信号。

[0023] 第二方面,本发明提供一种多路服务器的信号处理方法,应用于多路服务器中,所述多路服务器包括多个 CPU、多个南桥基板管理监控器输入输出 PBI 板,以及关键信号处理板,其中,一个 PBI 板监控至少两个 CPU;所述方法包括:

[0024] 所述关键信号处理板接收所述 PBI 板发送的关键信号;

[0025] 所述关键信号处理板根据接收到的所述关键信号向所述 PBI 板返回与所述关键信号相对应的回应信号,以使接收到所述回应信号的所述 PBI 板将所述回应信号发送给所述 PBI 板监控的 CPU。

[0026] 结合第二方面,在第二方面的第一种可能的实现方式中,所述关键信号处理板根据接收到的所述关键信号向 PBI 板返回与所述关键信号相对应的回应信号,包括:

[0027] 当所述多路服务器作为一个服务器进行上电时,所述关键信号处理板接收到所述多路服务器内的全部 PBI 板分别发送的第一类电源上电信号后,同时向全部 PBI 板发送作为与所述第一类电源上电信号相对应的回应信号的第二类电源上电信号,以使所述多路服务器内的全部 CPU 接收到第二类电源上电信号的时间差满足预设时间要求;

[0028] 若所述多路服务器内的全部 CPU 划分成至少两个子服务器,当任意一个子服务器上电时,所述关键信号处理板接收到所述子服务器内的全部 PBI 板分别发送的所述第三类电源上电信号后,同时向所述子服务器内的全部 PBI 板分别发送作为与所述第三类电源上电信号相对应的回应信号的第四类电源上电信号,以使所述子服务器内的全部 CPU 接收到所述第四类电源上电信号的时间差满足预设时间要求;

[0029] 其中,所述第一类电源上电信号和所述第三类电源上电信号均由所述 PBI 板检测到自身监控的各个 CPU 的电源均上电完成时产生。

[0030] 结合第二方面,在第二方面的第二种可能的实现方式中,所述关键信号处理板根据接收到的所述关键信号向 PBI 板返回与所述关键信号相对应的回应信号,包括:

[0031] 当所述多路服务器中插入新的 CPU 时,所述关键信号处理板接收到所述新的 CPU 对应的 PBI 板发送的第五类电源上电信号后,在预设时刻向所述新的 CPU 对应的 PBI 板发送作为与所述第五类电源上电信号相对应的回应信号的第六类电源上电信号,以使所述新的 CPU 接收到所述第六类电源上电信号的时间满足预设时间要求;

[0032] 所述第五类电源上电信号由所述 PBI 板检测到所述新的 CPU 上电完成时产生。

[0033] 结合第二方面的第二种可能的实现方式,在第二方面的第三种可能的实现方式中,所述关键信号处理板根据接收到的所述关键信号向 PBI 板返回与所述关键信号相对应的回应信号包括:

[0034] 若所述多路服务器作为一个服务器使用,所述关键信号处理板接收到任意一个 PBI 板发送的第一系统错误信号后,同时向所述多路服务器中除发送所述系统错误信号 PBI 之外的其它各个 PBI 板发送作为与所述第一系统错误信号对应的回应信号的第二系统错误信号;

[0035] 若所述多路服务器划分成至少两个子服务器,所述关键信号处理板获取所述多路服务器的子服务器信息,当接收到任意一个 PBI 板发送的第三系统错误信号时,根据所述子服务器信息确定与发送所述第三系统错误信号的 PBI 板属于同一个子服务器的其它各个 PBI 板;同时向所述子服务器中除发送所述第三系统错误信号的 PBI 板之外的其它各个 PBI 板发送作为与所述第三系统错误信号对应的回应信号的第四系统错误信号。

[0036] 结合第二方面的第三种可能的实现方式,在第二方面的第四种可能的实现方式中,所述方法还包括:

[0037] 当检测到当前 PBI 板发送的第一系统错误信号或第三系统错误由有效状态变为无效状态时,所述关键信号处理板接收下一个第一系统错误信号或第三系统错误信号。

[0038] 由以上技术方案可见,本发明实施例提供的多路服务器包括:多个 CPU、多个 PBI (PCHBMC IO,南桥基板管理监控器输入输出) 板和一个关键信号处理板,其中,BMC 为 Baseboard Management Controller,基板管理监控器。每个 PBI 板监控至少两个 CPU,且每个 PBI 板与自身所监控的 CPU 连接。各个 PBI 板的输出端分别连接所述关键信号处理板的输入端,各个 PBI 板的输入端分别连接所述关键信号处理板的输出端。由此可知,各个 CPU 分别与相应的 PBI 板连接,PBI 板直接与关键信号处理板连接。所述多路服务器通过关键信号处理板接收各个 PBI 板发出的关键信号,然后,同时向相应的 PBI 板发送对应的回应信号。各个 PBI 板直接与关键信号处理板连接,不需要切换开关进行串联,减少切换开关的数量,结构简单、可靠性高。而且减少了信号所经过的切换开关的数量,提高了信号质量。

附图说明

[0039] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,对于本领域普通技术人员而言,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0040] 图 1 为本发明实施例一种多路服务器的结构示意图;

[0041] 图 2 为本发明实施例一种 PBI 板的结构示意图;

[0042] 图 3 为本发明实施例一种多路服务器的信号处理方法的流程图。

具体实施方式

[0043] 为了使本技术领域的人员更好地理解本发明中的技术方案,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都应当属于本发明保护的范围。

[0044] 参见图 1,示出了本发明实施例提供的一种多路服务器的结构示意图,本实施例以 32 路服务器为例进行说明。当然,在本发明的其它实施例中,多路服务器可以是 16 路、8 路、4 路的服务器,PBI 板的数量可以根据实际需要灵活确定。

[0045] 如图 1 所示,所述多路服务器包括:32 个 CPU、至少 8 个 PBI 板和 1 个关键信号处理板。

[0046] 本实施例以 1 个 PBI 板监控 4 个 CPU 为例进行说明,例如,PBI1 监控 CPU1 ~ CPU4,PBI2 监控 CPU5 ~ CPU8,PBI3 监控 CPU9 ~ CPU12,PBI4 监控 CPU13 ~ CPU16,依次类推,PBI8 监控 CPU29 ~ CPU32。

[0047] 每个 PBI 板均连接自身监控的多个 CPU,CPU1 ~ CPU4 与 PBI1 连接,CPU5 ~ CPU8 与 PBI2 连接,依次类推,CPU29 ~ CPU32 与 PBI8 连接。

[0048] 由于 PBI 板上的 I/O 接口的数量有限,PBI 板通过自身的 I/O 接口连接自身监控的部分 CPU,自身监控的另一部分 CPU 通过背板连接 PBI 板。

[0049] 当然,在本发明的其它实施例中,可以是 1 个 PBI 板监控 2 个或者 8 个 CPU。如果 1 个 PBI 板监控 2 个 CPU,则 32 路服务器需要配置 16 个 PBI 板;如果 1 个 PBI 板监控 8 个 CPU,则 32 路服务器需要配置 4 个 PBI 板。

[0050] 各个 PBI 板均包括输入端和输出端,各个 PBI 板的输入端分别连接关键信号处理板 100 的输出端,各个 PBI 板的输出端分别连接关键信号处理板 100 的输入端。

[0051] PBI 板接收到 CPU 发送的关键信号或者自身监控 CPU 的状态而产生相应的关键信号后,直接将关键信号发送给关键信号处理板 100,由关键信号处理板 100 向相应的 PBI 板发送与所述关键信号相对应的回应信号。接收到回应信号的 PBI 板将所述回应信号发送给相应的 CPU,从而保证各个 CPU 接收到回应信号满足要求。

[0052] 其中,所述关键信号包括但不限于 Powergood 信号、Caterr 信号、PMSYNC 信号、TSC(Time Stamp Counte,时间戳计数器)信号、S4 信号、Thermtrip 信号、Powerbutton 信号、MSMI 信号等。

[0053] 不同的关键信号对应的回应信号也不同,不同的回应信号的功能也不同。总体而言,通过回应信号使多路服务器中的各个 CPU 之间的关键信号满足多路服务器的要求。例如,Powergood 信号的回应信号仍是 Powergood 信号,CPU 收到 Powergood 信号后启动。

[0054] 在多路服务器中,通过 Powergood 信号通知 CPU:外界的各个电源已经完成上电。多路服务器对 Powergood 信号有非常严格的时序要求,例如,在多路服务器开机时,所有 CPU 收到的 Powergood 信号的时间差必须监控在 10ns 以内,这样能最大限度提高 32 路服务器的系统性能。并记录向各个 CPU 发送 Powergood 信号的时刻为初始时刻 T0。

[0055] 例如,多路服务器作为一个服务器进行上电时,PBI 板检测到自身监控的 CPU 上电后,向关键信号处理板发送第一 Powergood 信号(即,第一类电源上电信号)。当关键信号处理板接收到服务器内的全部 PBI 板分别发送的第一类 Powergood 信号后,同时向多路服务器内全部的 PBI 板返回第二 Powergood 信号(即,第二类电源上电信号)。PBI 板接收到关键信号处理板发送的第二 Powergood 信号后,分别发送给自身监控的 CPU。这样,由关键信号处理板将第二 Powergood 信号统一发送给各个 PBI 板,然后,再由各个 PBI 板分别发送给自身监控的 CPU,从而保证多路服务器内的各个 CPU 接收到第二 Powergood 信号的时间差在预设时间(例如,10ns)内。

[0056] 在 CPU 实现热插拔、online(上线)或 offline(下线)时,新 online 的 CPU 接收到 Powergood 信号的时间必须是预设时刻,该预设时刻是 T0 加上预设周期,其中,不同厂商生产的 CPU,或者,同一厂商生产的不同型号的 CPU,规定的预设周期可能不相同。例如,Ivybridge 处理器设定的预设周期可以是 $N*864BCLKs$;Haswell 处理器设定的预设周期可以是 $N*384BCLKs$ 。其中,N 为正整数,且 1BCLK 是 10ns。

[0057] Caterr 信号表示系统发生了严重错误。

[0058] PMSYNC(Power Management Sync,电源管理同步)信号是节能控制信号。

[0059] S4 信号是多路服务器系统的一种状态信号,该信号由南桥芯片(PCH)发出,当 CPU 所在单板上的 CPLD(Complex Programmable Logic Device,复杂可编程逻辑器件)检测到 PCH 发出 S4 以后,控制整个多路服务器系统上电。

[0060] Thermtrip 信号是 CPU 过热时发出的信号。

[0061] Powerbutton 信号是触发服务器上的上电按钮时产生的信号,该信号用于通知 PCH 外界有按钮触发动作,需要上电开机或者下电关机。

[0062] MSMI 信号和 Caterr 信号功能完全一样,是最新的 Intel 平台增加的一个信号。

[0063] 本发明实施例提供的多路服务器包括多个 CPU、多个 PBI 板和一个关键信号处理

板。每个 PBI 板监控至少两个 CPU,且每个 PBI 板与自身所监控的 CPU 连接。各个 PBI 板的输出端分别连接所述关键信号处理板的输入端,各个 PBI 板的输入端分别连接所述关键信号处理板的输出端。关键信号处理板接收 PBI 板发送的关键信号后,向相应的所述 PBI 板返回与所述关键信号相对应的回应信号。接收到回应信号的 PBI 板将所述回应信号分别发送给自身监控的 CPU,以使 CPU 接收到的回应信号满足多路服务器的要求。由上述内容可知,所述多路服务器通过关键信号处理板接收各个 PBI 板发出的关键信号,然后向相应的 PBI 板发送相应的回应信号,能够保证各个 CPU 接收到回应信号满足多路服务器的要求。而且,各个 PBI 板直接与关键信号处理板连接,不需要切换开关进行串联,减少切换开关的数量,结构简单、可靠性高。而且,减少了信号所经过的切换开关的数量,提高了信号质量。

[0064] 此外,本发明实施例中各个 PBI 板之间没有通过转换开关串联,可以实现任意多个 PBI 组合构成一个硬分区,即灵活配置硬分区。例如,将 PBI1 和 PBI3 对应的 CPU 配置为一个硬分区,即配置为一个 8 路服务器;将 PBI2 和 PBI5 配置为一个硬分区,即配置为一个 8 路服务器。而现有技术中的多路服务器,由于 PBI 板之间通过转换开关串联,因此,在配置硬分区时只能将相邻的 PBI 配置为一个硬分区,不能实现灵活配置硬分区。

[0065] 图 2 示出了本发明实施例一种 PBI 板的内部结构示意图,如图 2 所示,每个 PBI 板均包括:转换开关 200、电平转换芯片 210、第一 CPLD220、南桥芯片 230。其中,转换开关 200 的数量与 PBI 板监控的 CPU 的数量一致,本实施例中每个 PBI 板内设置有四个转换开关。

[0066] 每个 CPU 的 I/O 接口通过转换开关 200 连接电平转换芯片 210 的第一 I/O 接口,电平转换芯片 210 的第二 I/O 接口连接第一 CPLD220 的第一 I/O 接口。

[0067] 第一 CPLD220 的第一输入端作为 PBI 板的输入端连接关键信号处理板 100 的输出端。第一 CPLD220 的第一输出端作为 PBI 板的输出端连接关键信号处理板 100 的输入端。第一 CPLD220 的第二输出端连接南桥芯片 230。

[0068] CPU 输出的信号提供给电平转换芯片 210 转换成 CPLD 能处理的高电平的信号,转换成高电平信号后更适合远距离传输。

[0069] 电平转换芯片 210 将转换后的信号发送给第一 CPLD220。第一 CPLD220 将接收到的信号发送给关键信号处理板 100。

[0070] 关键信号处理板 100 向相应的 PBI 内的第一 CPLD220 返回与接收到的关键信号相对应的回应信号。然后,由第一 CPLD220 控制相应的转换开关 200 闭合,将第一 CPLD220 接收到关键信号处理板 100 发送的信号发送给相应的 CPU。

[0071] 下面结合图 1 和图 2,以不同的关键信号为例介绍本发明实施例提供的多路服务器的关键信号处理过程。

[0072] (1) 32 路服务器作为一个服务器上电时,Powergood 信号的传输过程:

[0073] 11) SMM(System Management Module,系统管理模块)板通知关键信号处理板该多路服务器是 32 路服务器。SMM 板用于获取多路服务器系统的分区信息。

[0074] 12) 各个 PBI 板及其监控的各个 CPU 分别上电;各个 PBI 板检测到自己监控的各个 CPU 的电源均上电完成后,产生第一 Powergood 信号并发送给关键信号处理板,表明该 PBI 板及其监控的 CPU 已经上电完成。

[0075] 在一个实施例中,每个 PBI 板可以产生一个第一 Powergood 信号,当检测到自身控制的全部 CPU 的电源都上电完成时产生。

[0076] 各个 PBI 板的电源不需要同步上电,关键信号处理板通过延时或去抖处理避免各个 PBI 板上电过程带来的不稳定状态,使系统的可靠性更高。延时是指在关键信号处理板接收到关键信号后先不处理,延时预设时长等关键信号完全稳定后再进行处理。预设时长可以根据上电稳定时间设定,当然也可以根据实际需求设定。

[0077] 需要说明的是,各个 PBI 板将第一 Powergood 信号发送给关键信号处理板后,PBI 板不做任何处理,直到接收到关键信号处理板发送的第二 Powergood 信号。

[0078] 13) 关键信号处理板接收到全部 PBI 板分别发送的第一 Powergood 信号后,在同一时刻分别向全部 PBI 板发送第二 Powergood 信号,并记录发送第二 Powergood 信号的时刻为初始时刻 T_0 。

[0079] 此外,由于不同的单板之间传输信号具有离散性,以及信号的传输线路不同延时也不同,要求关键信号处理板到各个 PBI 板之间的传输线路的长度相等,这样能够保证各个 PBI 板收到第二 Powergood 信号的时刻保持一致,进而能够满足系统对各个 CPU 收到第二 Powergood 信号的时序要求。

[0080] 需要说明的是,在 CPU 热插拔时,目前的 BIOS(Basic Input Output System,基本输入输出系统)只能做到两个 CPU 同时 online(上线)或 offline(下线),因此,在多路服务器中,需要将两个 CPU 划分成一组,例如,32 路服务器一共有 16 组。

[0081] 为了使多路服务器上电与 CPU 热插拔时,对 Powergood 信号的处理过程相兼容。多路服务器上电时将每两个 CPU 划分成一组,每个 PBI 板针对每一组 CPU 产生一个第一 Powergood 信号,并分别发送给关键信号处理板。关键信号处理板需要针对每一组 CPU 向 PBI 板返回一个第二 Powergood 信号,保证第二 Powergood 信号的数量与第一 Powergood 信号的数量相同。例如,每个 PBI 板监控 4 个 CPU,两个 CPU 为一组,则每个 PBI 板产生两个第一 Powergood 信号。而且,关键信号处理板需要同时向每个 PBI 板返回两个第二 Powergood 信号。

[0082] 14) PBI 板中的第一 CPLD 将关键信号处理板发送的第二 Powergood 信号分别发送给自身监控的各个 CPU。

[0083] 关键信号处理板可以同时向各个 PBI 板分别发送第二 Powergood 信号,然后,再由各个 PBI 板将所述第二 Powergood 信号分别发送给相应的 CPU,从而保证各个 CPU 收到第二 Powergood 信号的时刻满足时序要求(即,各个 CPU 收到第二 Powergood 信号的时间差在 10ns 以内)。

[0084] (2) 当将 32 路服务器划分成多个子服务器时,子服务器上电的过程与上述过程类似。其中,一个子服务器即一个硬分区,每个子服务器可以单独工作。例如,32 路服务器可以分为 4 个 8 路的子服务器,也可以分为 1 个 16 路的子服务器和 2 个 8 路的子服务器。

[0085] 子服务器的上电过程如下:

[0086] 21) SMM 板通知关键信号处理板该多路服务器的硬分区信息。硬分区信息可以包括多路服务器包含几个硬分区、每个硬分区包含哪些 CPU 等信息。

[0087] 22) 各个硬分区的 PBI 板及其监控的各个 CPU 分别上电;PBI 板检测到自己监控的各个 CPU 的电源均上电完成后,产生第一 Powergood 信号(即,第三类电源上电信号)并发送给关键信号处理板,表明该 PBI 板及其监控的 CPU 已经上电完成。

[0088] 23) 关键信号处理板接收到硬分区内的 PBI 板发送的第一 Powergood 信号之后,确

定该 PBI 板所在的硬分区。

[0089] 可以根据来自 SMM 板的硬分区信息确定发送所述第一 Powergood 信号的 PBI 板所在硬分区的信息,例如,确定发送第一 Powergood 信号的 PBI 板位于哪个硬分区内,且该硬分区内还包括哪些 PBI 板等信息。

[0090] 24) 关键信号处理板接收到所述硬分区内的全部 PBI 板发送的第一 Powergood 信号之后,在同一时刻向该硬分区内的各个 PBI 板分别发送第二 Powergood 信号(即,第四类电源上电信号),并记录发送第二 Powergood 信号的时刻为初始时刻 T_0 。

[0091] 25) PBI 板中的第一 CPLD 将关键信号处理板发送的第二 Powergood 信号分别发送给自身监控的各个 CPU。

[0092] 通过上述的过程,能够实现硬分区内的各个 CPU 收到第二 Powergood 信号的时间满足时序要求。而且,各个硬分区分别上电、分别控制,相互之间不存在耦合关系。

[0093] (3) 当 CPU online/offline 时, Powergood 信号的传输过程

[0094] 当 CPU online/offline 时,32 路服务器作为一个服务器或划分成多个子服务器 Powergood 信号的传输过程相同,具体过程如下:

[0095] 31) 需要 online 的 CPU 插入多路服务器系统。

[0096] 32) PBI 板检测到新 online 的 CPU 上电完成后,产生第一 Powergood 信号(即,第五类电源上电信号),并将第一 Powergood 信号发送给关键信号处理板。

[0097] 33) 关键信号处理板接收到 PBI 板发送的第一 Powergood 信号后,在预设时刻将第二 Powergood 信号(即,第六类电源上电信号)发送给所述 PBI 板。

[0098] 所述预设时刻为所述初始时刻 T_0 加上预设周期。其中,不同厂商生产的 CPU,或者,同一厂商生产的不同型号的 CPU,规定的预设周期可能不相同,例如, Ivybridge 处理器对应的预设周期为 $N*864\text{BCLKs}$,相应的 Ivybridge 处理器对应的预设时刻为 $T_0+N*864\text{BCLKs}$; Haswell 处理器对应的预设周期为 $N*384\text{BCLKs}$,相应的 Haswell 处理器对应的预设时刻为 $T_0+N*384\text{BCLKs}$,其中, N 为正整数。

[0099] 34) 由所述 PBI 板将接收到的第二 Powergood 信号发送给相应的 CPU。

[0100] 通过上述的过程实现 CPU online 时, online 的 CPU 接收到的第二 Powergood 信号满足预设时间要求。

[0101] 下面以系统错误信号(即, Caterr 信号)为例,介绍本发明实施例提供的多路服务器的工作过程。

[0102] (4) 若 32 路服务器作为一个服务器使用,则 Caterr 信号的传输过程如下:

[0103] 41) PBI 板接收到 CPU 发送的第一 Caterr 信号(即,第一系统错误信号)时,该 PBI 板将所述第一 Caterr 信号发送给关键信号处理板。

[0104] 42) 关键信号处理板接收到的第一 Caterr 信号后,向除发送所述第一 Caterr 信号的 PBI 板之外的其它各个 PBI 板分别发送第二 Caterr 信号(即,第二系统错误信号)。

[0105] 关键信号处理板接收到第一 Caterr 信号后,根据与 Caterr 信号对应的标志位确定是否响应该第一 Caterr 信号。例如,当与 Caterr 信号对应的标志位为“0”时,关键信号处理板不再响应其它 CPU 发来的 Caterr 信号;当与 Caterr 信号对应的标志位为“1”时,关键信号处理板能够响应其它 CPU 发来的 Caterr 信号。

[0106] 当关键信号处理板接收到第一个 Caterr 信号后,将与 Caterr 信号对应的标志位

修改成二进制数“0”。若第一个 Caterr 信号由低电平变为高电平时,表明故障排除,此时,将与 Caterr 信号对应的标志位修改成二进制数“1”。

[0107] 关键信号处理板响应 CPU 发来的 Caterr 信号是指关键信号处理板接收到 CPU 发送的 Caterr 信号之后,将 Caterr 信号转发给与该 CPU 处于同一硬分区的其它 CPU,即把与该 CPU 处于同一硬分区的其它 CPU 的 Caterr 信号线由高电平拉成低电平,从而屏蔽其它 CPU 发送的 Caterr 信号。如果发出 Caterr 信号的 CPU 能够将该错误自行修复,则该 Caterr 信号就会消失,系统恢复正常;如果 CPU 无法自行修复,系统将会出现挂死现象。

[0108] 43) 各个 PBI 板将接收到的第二 Caterr 信号分别发送给相应的 CPU。

[0109] (5) 若 32 路服务器划分成多个子服务器使用,每个子服务器即一个硬分区。Caterr 信号的传输过程如下:

[0110] 51) SMM 板通知关键信号处理板多路服务器的硬分区信息。

[0111] 52) PBI 板接收到 CPU 发送的第三 Caterr 信号(即,第三系统错误信号)时,该 PBI 板将所述第三 Caterr 信号发送给关键信号处理板。

[0112] 53) 关键信号处理板接收到第三 Caterr 信号后,根据所述硬分区信息确定该 PBI 板对应的硬分区。

[0113] 即关键信号处理板确定第三 Caterr 信号对应的回应信号的广播范围,换言之,确定将与第三 Caterr 信号对应的回应信号发送给哪些 PBI 板。

[0114] 54) 关键信号处理板将第四 Caterr 信号(即,第四系统错误信号)发送给所述 PBI 板所在硬分区中的其它 PBI 板。

[0115] 55) 所述硬分区中的其它 PBI 板接收到第四 Caterr 信号后,发送给自己监控的 CPU。

[0116] 利用本实施例提供的多路服务器,关键信号处理板接收到 CPU 发送的 Caterr 信号后,能够通过 PBI 发送给相应的 CPU,由于 PBI 板之间未通过转换开关串联,而是直接与关键信号处理板连接,因此,控制过程简单,而且大大减少了 Caterr 信号的广播时间。另外,减少了信号所经过的切换开关的数量,能够提高信号质量。

[0117] 相应于上述的多路服务器,本发明还提供了多路服务器的信号同步方法实施例。

[0118] 参见图 3,为本发明实施例提供的一种多路服务器的信号处理方法的流程示意图,该方法应用于多路服务器的关键信号处理板中。所述多路服务器包括多个 CPU、多个 PBI 板和关键信号处理板,其中,一个 PBI 板监控至少两个 CPU;每个 PBI 板均连接自身监控的多个 CPU;各个 PBI 板均包括输入端和输出端,各个 PBI 板的输入端分别连接关键信号处理板的输出端,各个 PBI 板的输出端分别连接关键信号处理板的输入端。

[0119] 如图 3 所示,该方法可以包括以下步骤:

[0120] S110,关键信号处理板接收 PBI 板发送的关键信号。

[0121] 所述关键信号由所述 PBI 板所监控的 CPU 发送给所述 PBI 板,或者,由 PBI 板检测到 CPU 的相应状态后产生,例如,当 PBI 板检测到自己监控的各个 CPU 的电源均上电完成后,产生第一 Powergood 信号;又如,当 CPU 发生故障后,产生第一 Caterr 信号,并将第一 Caterr 信号发送给 PBI 板,由 PBI 板发送给关键信号处理板。

[0122] 所述关键信号包括但不限于 Powergood 信号、Caterr 信号、PMSYNC 信号、TSC 信号、S4 信号、Thermtrip 信号、Powerbutton 信号、MSMI 信号。

[0123] 其中 Powergood 信号和 Caterr 信号的处理过程如上述过程所述,此处不再赘述。

[0124] S120,关键信号处理板根据接收到的所述关键信号向 PBI 板返回与所述关键信号相对应的回应信号,以使接收到所述回应信号的 PBI 板将所述回应信号发送给所述 PBI 板监控的 CPU。

[0125] 由上述实施例可见,多路服务器可以包括 32 个 CPU、8 个 PBI 板和一个关键信号处理板。每个 PBI 板监控多个 CPU;各个 PBI 板的输出端均连接所述关键信号处理板的输入端,各个 PBI 板的输入端连接所述关键信号处理板的输出端;关键信号处理板接收 PBI 板发送的关键信号,产生与所述关键信号相对应的回应信号,并将所述回应信号返回给相应的 PBI 板。接收到回应信号的 PBI 板将所述回应信号分发给相应的 CPU。由上述内容可知,各个 PBI 板直接与关键信号处理板连接,不需要切换开关进行串联,减少切换开关的数量,结构简单、可靠性高,而且减少了信号所经过的切换开关提高了信号质量。

[0126] 若关键信号为多路服务器上电时的电源上电信号 (Powergood 信号),则步骤 S120 可以包括以下两种情况:

[0127] 其一,多路服务器作为一个服务器上电

[0128] 所述关键信号处理板接收到所述多路服务器内的全部 PBI 板分别发送的第一类电源上电信号后,同时向全部 PBI 板发送作为与所述第一类电源上电信号相对应的回应信号的第二类电源上电信号,以使所述多路服务器内的全部 CPU 接收到第二类电源上电信号的时间差满足预设时间要求。

[0129] 其中,所述第一类电源上电信号由所述 PBI 板检测到自身监控的各个 CPU 的电源均上电完成时产生。

[0130] 其二,所述多路服务器划分成至少两个子服务器,任意一个子服务器上电的过程如下:

[0131] 所述关键信号处理板接收到所述子服务器内的全部 PBI 板分别发送的所述第三类电源上电信号后,同时向所述子服务器内的全部 PBI 板分别发送作为与所述第三类电源上电信号相对应的回应信号的第四类电源上电信号,以使所述子服务器内的全部 CPU 接收到所述第四类电源上电信号的时间差满足预设时间要求。

[0132] 其中,所述第三类电源上电信号均由所述 PBI 板检测到自身监控的各个 CPU 的电源均上电完成时产生。

[0133] 若关键信号为 CPU online/offline 时的电源上电信号 (即,Powergood 信号),步骤 S120 可以包括以下过程:

[0134] 当所述多路服务器中插入新的 CPU 时,所述关键信号处理板接收到所述新的 CPU 对应的 PBI 板发送的第五类电源上电信号后,并在预设时刻向所述新的 CPU 对应的 PBI 板发送作为与所述第五类电源上电信号相对应的回应信号的第六类电源上电信号,以使所述新的 CPU 接收到所述第六类电源上电信号的时间满足预设时间要求。

[0135] 所述第五类电源上电信号由所述 PBI 板检测到所述新的 CPU 上电完成时产生。

[0136] 若关键信号为系统错误信号 (即 Caterr 信号),则步骤 S120 可以包括以下两种情况:

[0137] 其一,若多路服务器作为一个服务器使用

[0138] 关键信号处理板接收到任意一个 PBI 板发送的第一系统错误信号后,同时向所述

多路服务器中除发送所述系统错误信号 PBI 之外的其它各个 PBI 板发送第二系统错误信号 ;所述第二系统错误信号是与所述第一系统错误信号相对应的所述回应信号

[0139] 其二,多路服务器划分成至少两个子服务器使用

[0140] 关键信号处理板获取所述多路服务器的子服务器信息,当接收到任意一个 PBI 板发送的第三系统错误信号时,根据所述子服务器信息确定与发送所述第三系统错误信号的 PBI 板属于同一个子服务器的其它各个 PBI 板 ;同时向所述子服务器中除发送所述第三系统错误信号的 PBI 板之外的其它各个 PBI 板发送第四系统错误信号 ;所述第四系统错误信号是与所述第三系统错误信号相对应的所述回应信号。

[0141] 可选地,关键信号为 Caterr 信号时,所述方法还包括 :

[0142] 当检测到当前 PBI 板发送的第一系统错误信号或第三系统错误由有效状态变为无效状态时,所述关键信号处理板接收下一个第一系统错误信号或第三系统错误信号。

[0143] 其中,系统错误信号的状态是否有效可以通过检测与系统错误信号对应的标志位获得,在一个具体的实施例中,若与系统错误信号对应的标志位为二进制数“0”时,表明系统错误信号有效 ;若与系统错误信号对应的标志位为二进制数“1”时,表明系统错误无效。当然,也可以是标志位为“1”时,表示系统错误信号有效 ;标志位为“0”时,表示系统错误信号无效。

[0144] 通过以上的方法实施例的描述,所属领域的技术人员可以清楚地了解到本发明可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:只读存储器(ROM)、随机存取存储器(RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0145] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于装置或系统实施例而言,由于其基本相似于方法实施例,所以描述得比较简单,相关之处参见方法实施例的部分说明即可。以上所描述的装置及系统实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0146] 本发明可以在由计算机执行的计算机可执行指令的一般上下文中描述,例如程序模块。一般地,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本发明,在这些分布式计算环境中,通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中,程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

[0147] 需要说明的是,在本文中,诸如“第一”和“第二”等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意

在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0148] 以上所述仅是本发明的具体实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

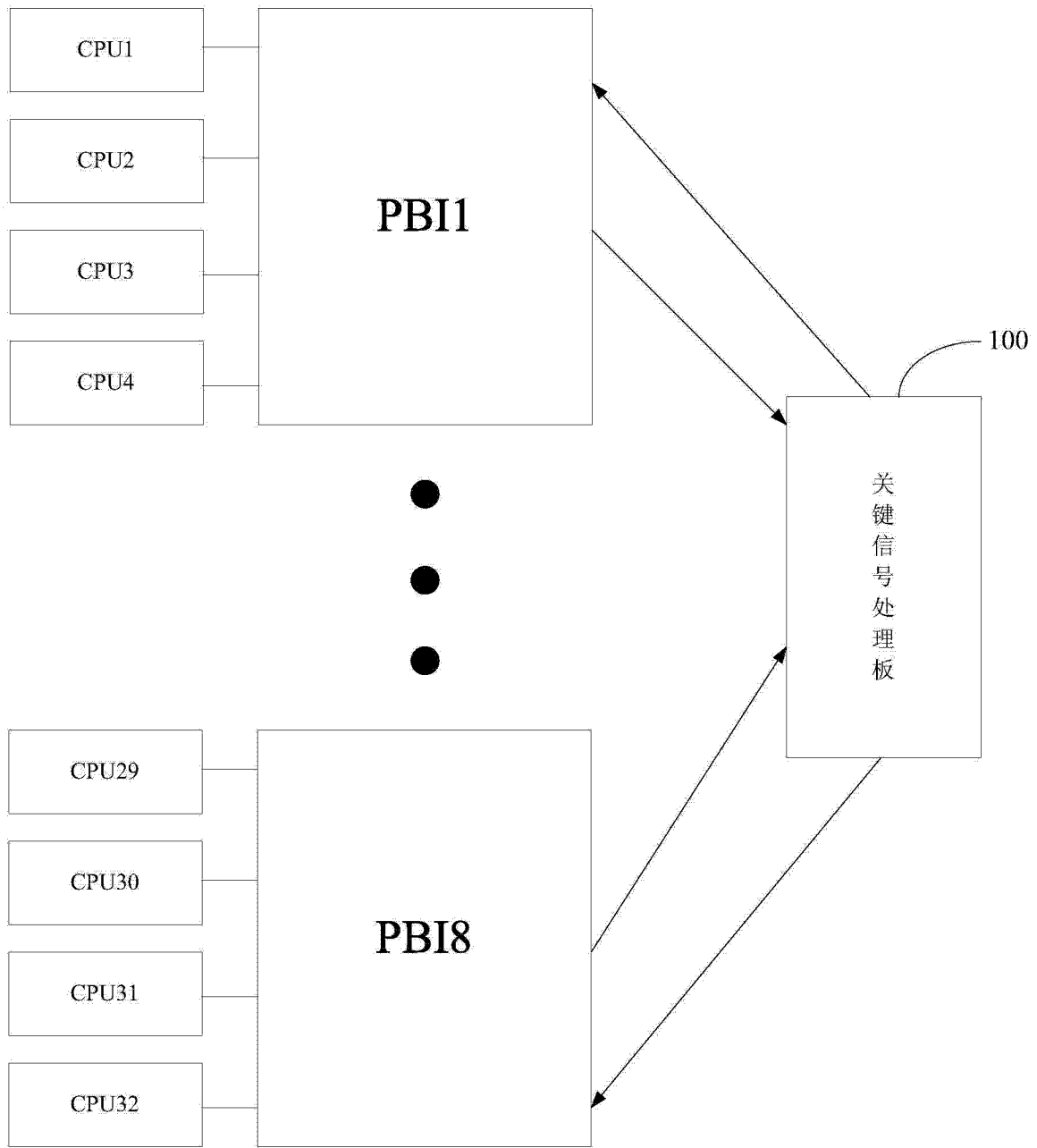


图 1

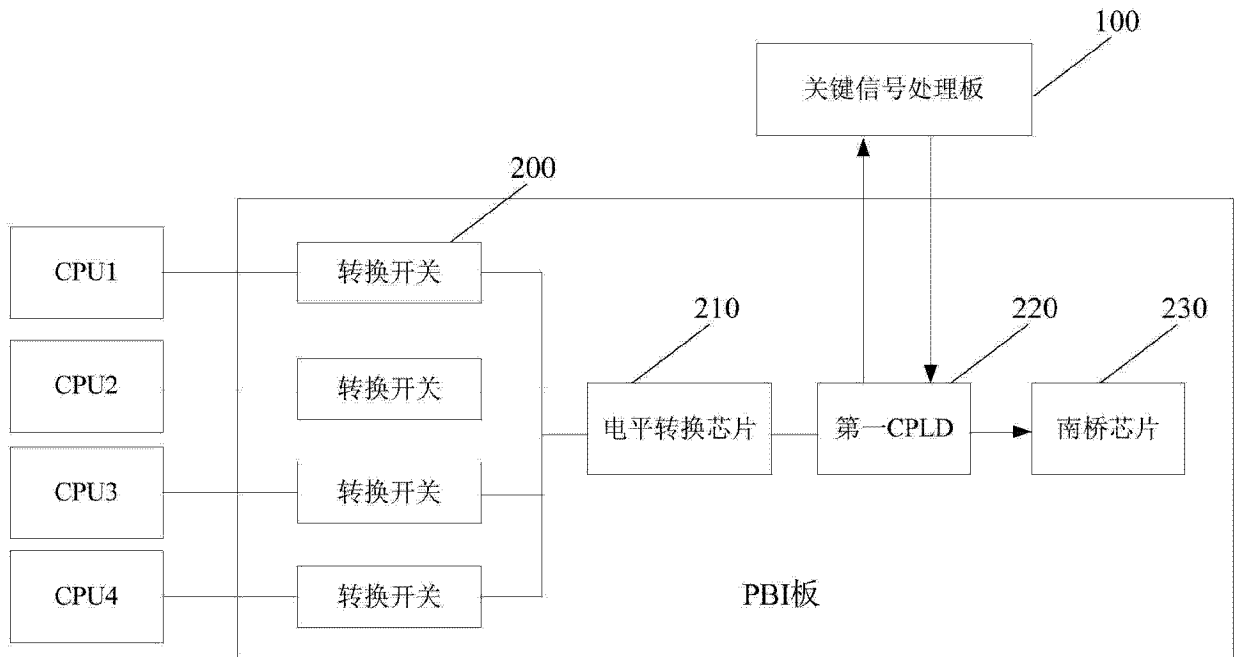


图 2

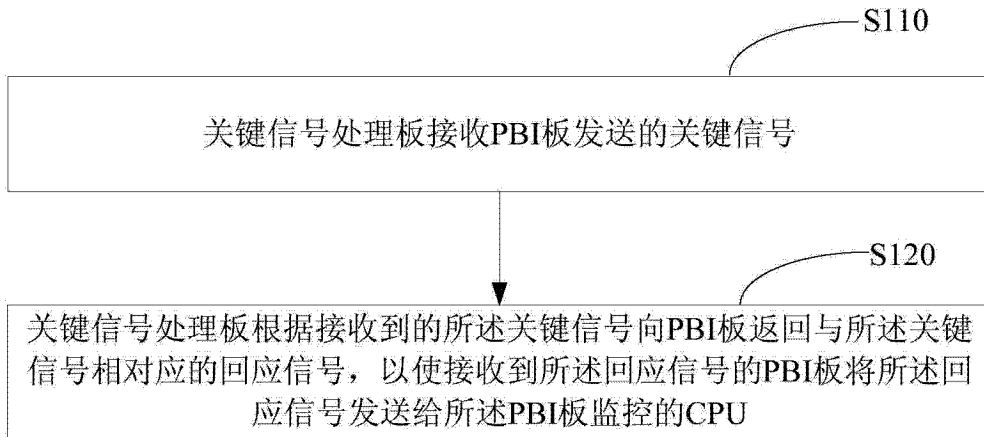


图 3